On asymptotic Lagrangian duality for nonsmooth optimization

Regina S. Burachik[1]

Received 5 March 2017; revised 11 November 2017

Abstract

Zero duality gap for nonconvex optimization problems requires the use of a generalized Lagrangian function in the definition of the dual problem. We analyze the situation in which the original problem is associated with a sequence of Lagrangian functions, which in turn defines a sequence of dual problems. Under a set of basic assumptions, we prove that the generated sequence of optimal dual values converges to the optimal primal value, and call the latter situation *strong duality* for the sequence of Lagrangian functions. As an application of our theory, we construct two sequences of augmented Lagrangians for general equality constrained optimization problems in finite dimensions which exhibit strong duality in this new sense.

DOI:10.21914/anziamj.v58i0.11874, Published 2017-11-23, as part of the Proceedings of the 18th Biennial Computational Techniques and Applications Conference. ISSN 1445-8810.

Contents

1 Introduction **C94**

2 Basic facts, definitions and assumptions **C97**
2.1 A limiting Lagrangian and its dual value C98
2.2 Asymptotic weak duality C99
2.3 Assumptions for asymptotic duality C102
2.4 Strong asymptotic duality C109

3 Equality constrained problems **C111**
3.1 When f_0 is not bounded below C115

4 Conclusions **C120**

References **C121**

1 Introduction

We consider the minimization problem (P) :

$$\text{minimize } f_0(x) \text{ subject to } x \text{ in } X_0$$

where X is a metric space, the subset $X_0 \subset X$ is closed, and the function $f_0 : X \to \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$ is lower semicontinuous. Denote by $d(\cdot, \cdot)$ the metric distance in X and let

$$M_P := \inf_{x \in X_0} f_0(x),$$

be the optimal value of the problem (P) .

We do not assume convexity nor differentiability on problem (P) . In this situation, a *duality approach* provides a convenient and elegant way of addressing the problem. Such an approach uses the information on problem (P) to construct a new optimization problem, called the *dual problem*. The objective function of the dual problem is defined by means of a *Lagrangian function*. In the present paper, a Lagrangian function is any function that reflects, in a way which will be made precise later on, the properties of problem (P). Since a Lagrangian function can be chosen in many different ways, many dual problems can be associated with the same problem (P).

A right choice of Lagrangian is the one which results in good duality properties. These properties are *zero duality gap*, which means that the primal problem (P) has the same optimal value as the dual problem, and *saddle point properties*, which means that a dual solution can provide a primal one. It is well known that this approach is very powerful when the problem is convex, and in this case the Lagrangian function is the classical (or ordinary) one. When the problem is nonconvex, then the classical approach no longer works and in general we will have a nonzero duality gap. In this case, a generalized Lagrangian function is needed. Indeed, it has been shown that for a certain family of *augmented Lagrangians*, both zero duality gap and saddle point properties hold [18, Chapter 11]. These duality properties have been extended to more general kinds of Lagrangians and more general frameworks (including infinite dimensional spaces) [11, 13, 17, 19, 20]. Many solution techniques for nonconvex optimization rely on the good duality properties of augmented Lagrangians [2, 3, 4, 5, 6, 7, 8, 9, 14]. Most of these papers share the following three features: (i) generate a primal-dual sequence at each iteration, (ii) under mild assumptions, prove that every accumulation point of the primal sequence is a solution of (P), and (iii) the dual problem is convex. Part (iii) makes the dual tractable and this allows one to devise algorithmic approaches which under mild assumptions produce a primal solution. The tractability of the dual problem, together with the good duality properties of the primal-dual sequence, justify a duality approach for solving (P), especially when the latter is nonconvex.

In spite of these benefits, the dual problem has some disadvantages, such as the lack of smoothness of the Lagrangian at points close to the solution set, even when the problem data is smooth. This motivated Huang et al. [17] to analyse suitable perturbations of a given Lagrangian function for nonlinear semidefinite programming problems. Moreover, if we use a duality approach for solving the primal problem it is convenient to incorporate current iterate information into the dual steps. This *dynamical update* allows one to incorporate current information and has the potential to produce accelerated versions of the duality scheme. Hence it makes sense to propose and analyse a theoretical framework that deals with a sequence of Lagrangians, and a sequence of dual problems. In the present paper, we consider the sequence of optimal dual values induced by a sequence of dual problems, and develop a general framework that ensures convergence of the optimal dual values to the optimal primal value.

First, we establish this desirable 'asymptotic' property under a basic set of assumptions (see Corollary 17). Second, we apply the theory to general equality constrained problems in finite dimensions. Namely, we construct two sequences of augmented Lagrangians for which strong duality (in the asymptotic sense) holds.

To our knowledge, only Huang et al. [17] and Burachik and Yang [10] consider a family of dual problems associated with a reference primal problem (P). Huang et al. study the case of the nonlinear semidefinite programming problem, and define a family of perturbations $\{\mathsf{L}_\varepsilon\}_\varepsilon$ of a fixed Lagrangian L. It is shown that the optimal dual values associated with L_ε converge to the optimal dual value induced by L when ε tends to zero. However, convergence to the optimal primal value requires rather strong assumptions on the problem. This motivated Burachik and Yang [10] to propose a more general analysis for the family of perturbations $\{\mathsf{L}_\varepsilon\}_\varepsilon$. Namely, using a set of basic assumptions, Proposition 2 and Lemma 2.1 in the paper by Burachik and Yang establish *strong asymptotic duality*, i.e.,

$$\liminf_{\varepsilon\downarrow 0} \mathsf{M}_{\mathsf{D}_\varepsilon} = \mathsf{M}_{\mathsf{P}}, \tag{1}$$

where $\mathsf{M}_{\mathsf{D}_\varepsilon}$ are the optimal values of the dual problems induced by L_ε. Burachik and Yang impose most of the hypothesis on the whole sequence of Lagrangians. Since it may not be simple to check the hypotheses for the whole sequence, the present paper develops an alternative analysis in which the hypotheses are imposed on a single function, namely the limit inferior of the sequence of Lagrangians. Having in mind a future convergence analysis of primal-dual techniques in which the Lagrangian function is updated at each iteration, we consider a sequence of Lagrangians $\{L_k\}_k$, instead of a family of perturbations $\{L_\varepsilon\}_\varepsilon$.

In Section 2 we give the basic definitions related to the asymptotic behaviour of the sequence of Lagrangians. We establish weak duality in the asymptotic sense (see Corollary 7). Strong asymptotic duality in the sense of (1) is established in Corollary 17. Section 3 is devoted to the prototypical example of a general equality constrained problem, and presents two results involving sequences of augmented Lagrangians. The first one establishes strong asymptotic duality when the objective function is bounded below, the augmenting function has a valley at zero, and the constraints satisfy a global error bound. The second result considers the case in which the objective function may be unbounded below, a global error bound holds for the constraints, and the augmenting function verifies an additional mild assumption. The last section, Section 4, contains some concluding remarks.

2 Basic facts, definitions and assumptions

In our analysis, assume that we have a sequence $\{L_k\}_{k\in\mathbb{N}}$ of Lagrangian functions. We denote the nonempty set of *dual variables* by Λ . Each $L_k : X \times \Lambda \to \mathbb{R}_\infty$ defines a dual function $q_k : \Lambda \to \mathbb{R}\cup\{-\infty\} := \mathbb{R}_{-\infty}$ as

$$q_k(\lambda) := \inf_{x\in X} L_k(x,\lambda), \tag{2}$$

and the associated dual problem (D_k) as

$$\max_{\lambda\in\Lambda} \; q_k(\lambda).$$

Let us denote as M_{D_k} the optimal value of (D_k) . The following definition relates the sequence of dual problems with problem (P) .

Definition 1. The sequence of Lagrangians $\{L_k\}_{k\in\mathbb{N}}$ exhibits *asymptotic duality* for (P) if

$$\liminf_{k\to\infty} M_{D_k} \geqslant M_P.$$

The sequence of Lagrangians $\{L_k\}_{k\in\mathbb{N}}$ exhibits *strong asymptotic duality* for (P) if

$$\liminf_{k\to\infty} M_{D_k} = M_P.$$

The sequence of Lagrangians $\{L_k\}_{k\in\mathbb{N}}$ exhibits *strong duality* for (P) if the $\lim_{k\to\infty} M_{D_k} \in \mathbb{R}$ and, moreover

$$\lim_{k\to\infty} M_{D_k} = M_P.$$

2.1 A limiting Lagrangian and its dual value

Our aim in this section is to establish weak duality results, both in terms of the sequence of optimal dual values, and in terms of the optimal value induced by the limit inferior of the sequence of Lagrangians.

Definition 2. For a given sequence $\{L_k\}$ of Lagrangians, the *Lagrangian induced by* $\{L_k\}$ is the function $L : X \times \Lambda \to \mathbb{R} \cup \{+\infty\} \cup \{-\infty\} =: \mathbb{R}_{\pm\infty}$ defined as

$$L(x, \lambda) := \liminf_{k\to\infty} L_k(x, \lambda), \quad \text{for all } x \in X,\ \lambda \in \Lambda. \tag{3}$$

The Lagrangian L defined in (3) induces the *dual function* $q : \Lambda \to \mathbb{R}_{-\infty}$ given by $q(\lambda) := \inf_{x\in X} L(x, \lambda)$ with *dual problem*

$$\sup_{\lambda\in\Lambda} q(\lambda), \tag{4}$$

and *optimal dual value* M_D .

2.2 Asymptotic weak duality

For deriving weak duality in the asymptotic sense, we need a Lagrangian which does not 'penalize' feasible points. Namely, we will assume that the Lagrangian verifies the following condition.

Definition 3. **(i)** $H_0(L)$ holds if

$$L(x,\lambda) \leqslant f_0(x), \text{ for all } x \in X_0, \lambda \in \Lambda. \tag{5}$$

For the sequence of Lagrangians, we relax the above inequality as follows.

(ii) $H_0(L_k)$ holds if there exists a sequence $0 \leqslant r_k \downarrow 0$ and $k_0 \in \mathbb{N}$ such that

$$L_k(x,\lambda) \leqslant f_0(x) + r_k, \text{ for all } k \geqslant k_0, \tag{6}$$

all $x \in X_0$ and all $\lambda \in \Lambda$.

Remark 4. Consider the relationship between assumptions $H_0(L)$ and $H_0(L_k)$. If $H_0(L)$ holds then we can show that for every sequence $0 \leqslant r_k \downarrow 0$ there exists an infinite set $J := \{p_0 < p_1 < p_2 < \ldots < p_k <\} \subset \mathbb{N}$ such that

$$L_{p_k}(x,\lambda) \leqslant f_0(x) + r_k, \tag{7}$$

for all $k \geqslant \mathbb{N}$. Indeed, since $r_1 > 0$ we have from $H_0(L)$ that

$$L(x,\lambda) = \liminf_{k\to\infty} L_k(x,\lambda) = \sup_n \inf_{k\geqslant n} L_k(x,\lambda) \leqslant f_0(x) < f_0(x) + r_1.$$

So for every $n \in \mathbb{N}$ we have that

$$\inf_{k\geqslant n} L_k(x,\lambda) < f_0(x) + r_1. \tag{8}$$

Take $n := 1$ in (8) to find $p_1 \geqslant 1$ such that

$$L_{p_1}(x,\lambda) < f_0(x) + r_1.$$

Now using (8) with $n := p_1 + 1$ and $r_2 > 0$ instead of r_1, gives a $p_2 > p_1$ such that $L_{p_2}(x, \lambda) < f_0(x) + r_2$. We can thus define inductively a set J as in (7). Therefore (6) holds for a subsequence of $\{L_k\}$. Now let us show that condition $H_0(L_k)$ implies $H_0(L)$. Indeed, by definition of L we can write

$$L(x, \lambda) = \liminf_k L_k(x, \lambda) \leqslant f_0(x) + \liminf_k r_k = f_0(x),$$

for all $x \in X_0$, $\lambda \in \Lambda$.

Our next step is to establish weak duality in the asymptotic sense. Part (i) of the next proposition follows similar steps as those in the paper by Burachik and Yang [10, Lemma 2.1] and Theorem 4, and hence we omit it here. However, part (ii) of the proposition is new.

Proposition 5. Let L and $\{L_k\}$ be as in Definition 2. The following hold.

(i) If $H_0(L)$ holds, then we have

$$M_D \leqslant M_P,$$

and for every sequence $0 \leqslant r_k \downarrow 0$ there exists $J := \{p_0 < p_1 < p_2 < \ldots < p_k <\} \subset \mathbb{N}$ such that

$$M_{D_{p_k}} \leqslant M_P + r_k, \text{ for all } k \geqslant \mathbb{N}.$$

In particular, we have

$$\liminf_k M_{D_k} \leqslant M_P.$$

(ii) Moreover, if we have that for some $k_0 \in \mathbb{N}$, $H_0(L_k)$ holds for all $k \geqslant k_0$, then

$$\limsup_k M_{D_k} \leqslant M_P.$$

Proof: For the proof of part (i), see [10, Lemma 2.1] and Theorem 4. (ii) For $k \geqslant k_0$ we can write

$$\begin{aligned} M_{D_k} &= \sup_{\lambda \in \Lambda} q_k(\lambda) \\ &= \sup_{\lambda \in \Lambda} \inf_{x \in X} L_k(x, \lambda) \\ &\leqslant \sup_{\lambda \in \Lambda} \inf_{x \in X_0} L_k(x, \lambda) \\ &= r_k + \inf_{x \in X_0} f_0(x) \\ &= r_k + M_P, \end{aligned}$$

which gives

$$\limsup_k M_{D_k} \leqslant M_P + \limsup_k r_k = M_P.$$

♠

Remark 6. In our context, it is natural to assume that $M_P < \infty$. In this situation, if $H_0(L)$ holds, then Proposition 5 yields $M_D < \infty$ and $M_{D_k} < \infty$ for k in an infinite subset of $\mathbb{N}$. Moreover, if $H_0(L_k)$ holds for k large enough, then there exists k_1 such that $M_{D_k} < \infty$ for $k \geqslant k_1$.

The following corollary, which is new, shows how asymptotic duality implies weak duality.

Corollary 7. Let L and $\{L_k\}$ be as in Definition 2. The following hold.

(i) If $H_0(L)$ holds and the sequence of Lagrangians exhibits asymptotic duality for (P), then it verifies that

$$M_D \leqslant \liminf_k M_{D_k} = M_P.$$

(ii) Moreover, if we have that for some $k_0 \in \mathbb{N}$, $H_0(L_k)$ holds for all L_k with $k \geqslant k_0$, then

$$M_D \leqslant \lim_k M_{D_k} = M_P.$$

Proof: (i) From Proposition 5 and the asymptotic duality assumption,

$$\liminf_k M_{D_k} \leqslant M_P \leqslant \liminf_k M_{D_k},$$

which readily gives the equality. The inequality follows from Proposition 5. Statement (ii) follows from the second part of Proposition 5, which yields

$$\limsup_k M_{D_k} \leqslant M_P \leqslant \liminf_k M_{D_k},$$

which readily implies the existence of the limit. ♠

2.3 Assumptions for asymptotic duality

The previous section establishes weak duality under the assumption of asymptotic duality. In this section we study the conditions under which asymptotic duality holds. Namely, we consider an assumption (see H_1 below), which turns out to be not only sufficient but also necessary for asymptotic duality to hold. A version of this assumption for a single Lagrangian was first introduced by Burachik and Rubinov [12], while the version for a family of Lagrangians was considered by Burachik and Yang [10].

In contrast with the analysis of the previous section, which considers assumption H_0 describing how our Lagrangian functions should behave over the constraint set X_0, we now impose an assumption that takes care of the behaviour of the Lagrangians outside the constraint set.

In what follows we will always assume that $M_P < \infty$ and fix $\delta > 0$. We will use the set

$$X(\delta) := \{x \in X \,:\, d(x, X_0) < \delta\}.$$

Definition 8. Let $\Lambda_0 \subset \Lambda$, then

(i) $H_1(\Lambda_0)$ *holds for* L if for all $\alpha < M_P$ and all $\delta > 0$, the inequality

$$\sup_{\lambda\in\Lambda_0}\left[\inf_{x\notin X(\delta)} L(x,\lambda)\right] > \alpha,$$

holds. We refer to this as $H_1(\Lambda_0, L)$ holds.

(ii) $H_1(\Lambda_0)$ *holds for* $\{L_k\}$ if for all $\alpha < M_P$ and for all $\delta > 0$, there exists $k_0 \in \mathbb{N}$ such that the inequality

$$\sup_{\lambda\in\Lambda_0}\left[\inf_{x\notin X(\delta)} L_k(x,\lambda)\right] > \alpha,$$

holds for all $k \geqslant k_0$. We refer to this as $H_1(\Lambda_0, L_k)$ holds for k large enough.

Remark 9. Assumption H_1 may appear artificial. In order to justify its use, we prove that it is in fact necessary for zero duality gap to hold. Moreover, it is also necessary for the family $\{L_k\}$ to exhibit asymptotic duality. The proof of part (i) in the proposition below follows similar steps to those of Burachik and Rubinov [12, Theorem 2.1], which has stronger assumptions than those used here. Namely, it assumes zero duality gap and the condition H_0. Since statement (i) is still true without these assumptions, we include its proof here. Regarding part (ii), this follows similar steps as those of Burachik and Yang [10, Proposition 1], and thus we omit its proof here.

Proposition 10. *Assume there exists* $\Lambda_0 \subset \Lambda$ *such that*

$$M_D = \sup_{\lambda\in\Lambda_0} q(\lambda) = \sup_{\lambda\in\Lambda_0} q_k(\lambda), \tag{9}$$

for all $k \in \mathbb{N}$. *The following hold:*

(i) If $M_P \leqslant M_D$ *then we must have that* $H_1(\Lambda_0)$ *holds for* L.

(ii) If the sequence $\{L_k\}$ *exhibits asymptotic duality for* (P), *then we must have that* $H_1(\Lambda_0)$ *holds for* $\{L_k\}$.

Proof: (i) Assume that $M_P \leqslant M_D$. We claim that $H_1(\Lambda_0)$ holds for L. Indeed, for any $\delta > 0$, let $Q(\lambda,\delta) := \inf_{\{x\in X\,:\,d(x,X_0)\geqslant\delta\}} L(x,\lambda)$. We can write

$$\begin{aligned} M_P &\leqslant M_D \\ &= \sup_{\lambda\in\Lambda_0} q(\lambda) \\ &= \sup_{\lambda\in\Lambda_0} \inf_{x\in X} L(x,\lambda) \\ &\leqslant \sup_{\lambda\in\Lambda_0} \inf_{\{x\in X\,:\,d(x,X_0)\geqslant\delta\}} L(x,\lambda) \\ &= \sup_{\lambda\in\Lambda_0} Q(\lambda,\delta), \end{aligned} \tag{10}$$

where we used the definition of M_D in the first equality, the definition of q in the second equality, and the definition of Q in the last equality. If $H_1(\Lambda_0)$ does not hold for L, there exist $\alpha_0 < M_P$ and $\delta_0 > 0$ such that

$$\sup_{\lambda\in\Lambda_0} Q(\lambda,\delta_0) \leqslant \alpha_0.$$

Combining this expression with (10) for $\delta := \delta_0$ we have

$$M_P \leqslant \sup_{\lambda\in\Lambda_0} Q(\lambda,\delta_0) \leqslant \alpha_0,$$

which contradicts the assumption that $M_P > \alpha_0$. Hence the claim is true and $H_1(\Lambda_0)$ holds for L. For the proof of part (ii) see the paper by Burachik and Yang [10, Proposition 1]. ♠

Remark 11. If for every fixed $\lambda \in \Lambda_0$ the functions $L_k(\cdot,\lambda)$ are minorized by $L(\cdot,\lambda)$ on $X\backslash X_0$, then it directly follows from the definitions that $H_1(\Lambda_0, L)$ implies $H_1(\Lambda_0, L_k)$ for k large enough. Namely, assume that

(i) $H_1(\Lambda_0, L)$ holds.

(ii) There exists $k_0 \in \mathbb{N}$ such that

$$L(x,\lambda) \leqslant L_k(x,\lambda), \quad \text{for all } k \geqslant k_0,$$

for every $x \notin X_0$, $\lambda \in \Lambda_0$.

In this situation, $H_1(\Lambda_0, L_k)$ holds for $k \geqslant k_0$. Indeed, for all $\alpha < M_P$ and all $\delta > 0$, use (ii) to write, for $k \geqslant k_0$,

$$\sup_{\lambda \in \Lambda_0} \left[\inf_{x \notin X(\delta)} L_k(x, \lambda) \right] \geqslant \sup_{\lambda \in \Lambda_0} \left[\inf_{x \notin X(\delta)} L(x, \lambda) \right] > \alpha.$$

We now consider the following level-boundedness assumption on L.

$H_2(\Lambda_0, L)$: For all $\alpha < M_P$ there exists $\hat{r} > 0$ and $\hat{\lambda} \in \Lambda_0$ such that the set

$$\{x \in X(\hat{r}) \ : \ L(x, \hat{\lambda}) \leqslant \alpha\},$$

is compact.

Our next assumption is taken from [10].

$H_3(\Lambda_0, L)$: There exists $\Lambda_0 \subset \Lambda$ such that

$$f_0(x) \leqslant L(x, \lambda), \tag{11}$$

for all $x \in X$, $\lambda \in \Lambda_0$.

Remark 12. If the inequality in (11) is strict, then, from the definition of L, we can find $l_0 \in \mathbb{N}$ such that $H_3(\Lambda_0, L_k)$ holds for all $k \geqslant l_0$. Otherwise, the most we can say is that for every $\varepsilon > 0$ there exists $k_0 \in \mathbb{N}$ such that

$$f_0(x) \leqslant L_k(x, \lambda) + \varepsilon,$$

for every $k \geqslant k_0$. On the other hand, if $H_3(\Lambda_0, L_k)$ holds for k large enough, then (11) holds.

Lemma 13. Assume that f_0 is lower semicontinuous, and that $H_2(\Lambda_0, L)$ and $H_3(\Lambda_0, L)$ hold. Then for every $\alpha < M_P$ there exists $r > 0$ and $\lambda \in \Lambda_0$ such that

$$X(r) = \{x \in X \ : \ d(x, X_0) < r\} \subset \{x \in X \ : \ L(x, \lambda) > \alpha\}. \tag{12}$$

Proof: Assume that for some $\alpha_1 < M_P$ the inclusion is false for all $r > 0$ and all $\lambda \in \Lambda_0$. For this given $\alpha_1 < M_P$, take $\hat{r} > 0$ and $\hat{\lambda} \in \Lambda_0$ as in $H_2(\Lambda_0, L)$. In particular, the inclusion (12) will be false for $r_n := \hat{r}/n > 0$ and $\lambda := \hat{\lambda}$. This implies the existence of a sequence $\{x_n\}$ such that

$$d(x_n, X_0) \leqslant \hat{r}/n \leqslant \hat{r}, \text{ and } f_0(x_n) \leqslant L(x_n, \hat{\lambda}) \leqslant \alpha_1, \tag{13}$$

where we used $H_3(\Lambda_0, L)$ in the right-most expression. By (13) and H_2, the sequence $\{x_n\}$ has an accumulation point $\bar{x}$, which must belong to X_0 because X_0 is closed and $d(x_n, X_0)$ tends to zero. Using (13) and the lower semicontinuity of f_0 we obtain $f_0(\bar{x}) \leqslant \alpha_1 < M_P$, contradicting the fact that $\bar{x} \in X_0$. Hence inclusion (12) must hold for some $r > 0$ and some $\lambda \in \Lambda_0$. ♠

Our next step is to prove that, under assumptions H_1–H_3, the sequence $\{L_k\}$ exhibits asymptotic duality for (P). The next proposition extends one of Burachik and Yang [10, Proposition 2]. Indeed, our assumption H_2 is less restrictive that the one used by Burachik and Yang. Moreover, since the latter is proved for a family $\{L_\varepsilon\}_{\varepsilon>0}$ Lagrangians, we adapt the proof to our case. To reflect the penalty properties of the Lagrangian, and the fact that the dual variables should be unbounded above, we introduce a special (albeit natural) assumption on the structure of the set Λ_0, which is stated as follows. Our examples in Section 3 will all verify this natural assumption (see Lemma 22(ii) and Lemma 28(ii)).

Definition 14. We say that Λ_0 is *directed for* L if for any pair of dual variables $\lambda_1, \lambda_2 \in \Lambda_0$ there exists $\lambda_3 \in \Lambda_0$ such that

$$\max\{L(x, \lambda_1), L(x, \lambda_2)\} \leqslant L(x, \lambda_3), \quad \text{for all } x \in X.$$

Remark 15. As an illustration of Definition 14, we recall the augmented Lagrangian given by Rockafellar and Wets [18, Definition 11.55]. Consider an equality constrained problem (P) in which the constraint set $X_0 := \{x \in \mathbb{R}^n : h(x) = 0\}$ is defined using a continuous function $h : \mathbb{R}^n \to \mathbb{R}^m$.

Consider also an *augmenting function* $\sigma : \mathbb{R}^m \to \mathbb{R}_\infty$ such that $\min \sigma = 0$ and $\operatorname{argmin} \sigma = \{0\}$. For this case the set of dual variables is $\Lambda := \mathbb{R}^m \times \mathbb{R}_+$. With a suitable duality parametrization, the *augmented Lagrangian* $L : \mathbb{R}^n \times [\mathbb{R}^m \times \mathbb{R}_+] \to \mathbb{R}_{\pm\infty}$ becomes

$$L(x, (u, c)) := f_0(x) + c\sigma(h(x)) + \langle u, h(x)\rangle,$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbb{R}^m$. The set $\Lambda_0 := \{0\} \times \mathbb{N} \subset \Lambda$ is clearly *directed for* L. Indeed, given $(0, p), (0, q) \in \{0\} \times \mathbb{N}$ then let $r := \max\{p, q\}$ so we have

$$\begin{aligned}
\max\{L(x, (0, p)), L(x, (0, q))\} &= \max\{f_0(x) + p\sigma(h(x)), f_0(x) + q\sigma(h(x))\} \\
&= f_0(x) + \sigma(h(x)) \max\{p, q\} \\
&= f_0(x) + r\sigma(h(x)) \\
&= L(x, (0, r)),
\end{aligned}$$

and the property of Definition 14 is satisfied.

Proposition 16. *Consider a sequence $\{L_k\}$ and L as in Definition 2. Let Λ_0 be directed for L, as in Definition 14. Assume further that*

(a) f_0 *is lower semicontinuous,*

(b) $H_1(\Lambda_0, L)$ *holds, and there exists* l_0 *such that* $L(x, \lambda) \leqslant L_k(x, \lambda)$ *for all* $k \geqslant l_0$ *and all* $\lambda \in \Lambda_0$,

(c) $H_2(\Lambda_0, L)$ *holds, and*

(d) $H_3(\Lambda_0, L)$ *holds.*

Then asymptotic duality holds for $\{L_k\}$. Namely,

$$M_P \leqslant \liminf_{k\to\infty} M_{D_k}.$$

Proof: Assume the conclusion is not true. In this case there exists α_1 such that

$$\liminf_{k\to\infty} M_{D_k} < \alpha_1 < M_P.$$

This yields the existence of a subsequence $\{k_j\}_{j\in\mathbb{N}}$ such that for all $j \in \mathbb{N}$ and

$$M_{D_{k_j}} < \alpha_1 < M_P. \tag{14}$$

By (a), (c), (d), and Lemma 13 for $\alpha := \alpha_1$, there exists $r_0 > 0$ and $\lambda_0 \in \Lambda_0$ such that

$$X(r_0) = \{x \in X \,:\, d(x, X_0) < r_0\} \subset \{x \in X \,:\, L(x, \lambda_0) > \alpha_1\}. \tag{15}$$

Using (b) for $\alpha := \alpha_1 < M_P$ and $\delta =: r_0 > 0$, we obtain the existence of $\lambda_1 \in \Lambda_0$ such that

$$\inf_{x \notin X(r_0)} L(x, \lambda_1) > \alpha_1. \tag{16}$$

Using the directed property of Λ_0, there exists $\lambda_2 \in \Lambda_0$ such that

$$\max\{L(x, \lambda_0), L(x, \lambda_1)\} \leqslant L(x, \lambda_2), \quad \text{for all } x \in X.$$

So we have

$$\begin{aligned} X(r_0) &= \{x \in X \,:\, d(x, X_0) < r_0\} \\ &\subset \{x \in X \,:\, L(x, \lambda_0) > \alpha_1\} \\ &\subset \{x \in X \,:\, L(x, \lambda_2) > \alpha_1\}. \end{aligned} \tag{17}$$

We also deduce from (16), the definition of λ_2 and (b) that

$$\inf_{x \notin X(r_0)} L_{k_j}(x, \lambda_2) \geqslant \inf_{x \notin X(r_0)} L(x, \lambda_2) \geqslant \inf_{x \notin X(r_0)} L(x, \lambda_1) > \alpha_1, \tag{18}$$

for $k_j \geqslant l_0$. Assume that $k_j \geqslant l_0$, then using (17) and (b) we have

$$q_{k_j}(\lambda_2) \leqslant M_{D_{k_j}} < \alpha_1 \leqslant \inf_{x \in X(r_0)} L(x, \lambda_2) \leqslant \inf_{x \in X(r_0)} L_{k_j}(x, \lambda_2). \tag{19}$$

The above expressions implies that

$$M_{D_{k_j}} \geqslant q_{k_j}(\lambda_2) = \inf_{x \notin X(r_0)} L_{k_j}(x, \lambda_2) > \alpha_1,$$

because (19) shows that the infimum value $q_{k_j}(\lambda_2)$ cannot be attained over the set $X(r_0)$. The above expression contradicts (14), and hence asymptotic duality holds. ♠

2.4 Strong asymptotic duality

Here we study conditions that ensure

$$\liminf_{k} M_{D_k} = M_D = M_P. \tag{20}$$

Corollary 17. *Consider a sequence* $\{L_k\}$ *and* L *as in Definition 2. Assume that all hypotheses of Proposition 16 hold. Then*

(i) If we also have that $H_0(L)$ *holds, then*

$$M_D = \liminf_{k\to\infty} M_{D_k} = M_P.$$

(ii) Moreover, if we have that for some $k_0 \in \mathbb{N}$, $H_0(L_k)$ *holds for all* $k \geqslant k_0$, *then*

$$\lim_{k\to\infty} M_{D_k} = M_D = M_P.$$

Proof: By Proposition 16, asymptotic duality holds for $\{L_k\}$. Combine this fact with Corollary 7(i) and assumption H_0 to conclude that $M_D \leqslant \liminf_{k\to\infty} M_{D_k} = M_P$. Recall that

$$M_D = \sup_{\lambda\in\Lambda} q(\lambda),$$

where $L(x, \lambda) = \liminf_{k\to\infty} L_k(x, \lambda)$. To complete the proof, we need to prove that $M_D \geqslant M_P$. Assume on the contrary that we have $M_D < \gamma < M_P$ for some γ. By Lemma 13, there exists $r_0 > 0$ and $\lambda_0 \in \Lambda_0$ such that

$$X(r_0) \subset \{x : L(x, \lambda_0) > \gamma\}. \tag{21}$$

Now using assumption $H_1(\Lambda_0, L)$ with $\delta := r_0$ and $\alpha := \gamma$ provides a $\lambda_1 \in \Lambda_0$ such that

$$\inf_{x\notin X(r_0)} L(x, \lambda_1) > \gamma. \tag{22}$$

As in the proof of Proposition 16, we use the directed property of Λ_0 to find $\lambda_2 \in \Lambda_0$ such that

$$\max\{L(x, \lambda_0), L(x, \lambda_1)\} \leqslant L(x, \lambda_2), \quad \text{for all } x \in X.$$

Then we have

$$\begin{aligned} X(r_0) &= \{x \in X : d(x, X_0) < r_0\} \\ &\subset \{x \in X : L(x, \lambda_0) > \gamma\} \\ &\subset \{x \in X : L(x, \lambda_2) > \gamma\}. \end{aligned} \tag{23}$$

From (22) and the definition of λ_2 we have

$$\inf_{x \notin X(r_0)} L(x, \lambda_2) \geqslant \inf_{x \notin X(r_0)} L(x, \lambda_1) > \gamma. \tag{24}$$

Using (24) and (23) we can write

$$q(\lambda_2) \leqslant M_D < \gamma \leqslant \inf_{x \in X(r_0)} L(x, \lambda_2). \tag{25}$$

The above expression implies that the infimum value of $q(\lambda_2)$ cannot be attained at points in $X(r_0)$, namely,

$$M_D \geqslant q(\lambda_2) = \inf_{x \notin X(r_0)} L(x, \lambda_2) > \gamma,$$

where we also used (24) in the last inequality. The above expression contradicts the assumption on M_D and hence we must have $M_D \geqslant M_P$. This establishes part (i). As for (ii), we apply Corollary 7(ii). Indeed, we know that we have asymptotic duality. This fact, together with assumption $H_0(L_k)$ holding for all $k \geqslant k_0$, implies that $\lim_{k\to\infty} M_{D_k} = M_P$. Since we already have $M_D = M_P$, the proof is complete. ♠

3 Equality constrained problems

In this section we construct sequences of augmented Lagrangians that verify strong asymptotic duality for equality constrained problems. To establish our claims we will apply the results of the previous section.

Consider the instance of problem (P) in which $X = \mathbb{R}^n$ with a norm $\|\cdot\|$. Let $X_0 := \{x \in \mathbb{R}^n \mid h(x) = 0\}$, with $h : \mathbb{R}^n \to \mathbb{R}^m$ a continuous function and $f_0 : \mathbb{R}^n \to \mathbb{R}_\infty$ a lower semicontinuous function. Let $\sigma : \mathbb{R}^m \to \mathbb{R}_\infty$ be such that

$$\sigma(z) \geqslant 0 \text{ for all } z \in \mathbb{R}^m \text{ and } \sigma(z) = 0 \text{ if and only if } z = 0. \tag{26}$$

The following additional assumption on σ is often used in the context of augmented Lagrangians [11, 20].

Definition 18. We say that σ *has a valley at* 0 if for every $\delta > 0$ we have $c_\delta = \inf_{\|y\| \geqslant \delta} \sigma(y) > 0$.

The following is a well-known assumption in the literature [16].

Definition 19. We say that a *global error bound* holds for a function $h : \mathbb{R}^n \to \mathbb{R}^m$ if there exists $\gamma > 0$ such that

$$d(x, X_0) \leqslant \gamma \|h(x)\|,$$

for all $x \in \mathbb{R}^n$.

Remark 20. For linear programming problems, a global error bound as in Definition 19 holds by *Hoffman's lemma* [15, Theorem 11.26].

We consider the following sequence of augmented Lagrangians.

Definition 21. Take $\Lambda := \mathbb{R}^m \times \mathbb{R}_+$. For each $k \in \mathbb{N}$, let $F_k : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be such that

$$F_k(0, z) = F_k(z, 0) = 0, \tag{27}$$

for all $z \in \mathbb{R}^m$. For the sequence $\{F_k\}_{k\in\mathbb{N}}$ and $\lambda = (u, c) \in \mathbb{R}^m \times \mathbb{R}_+$, we construct the sequence $L_k : \mathbb{R}^n \times \Lambda \to \mathbb{R}_\infty$ as follows

$$L_k(x, \lambda) := f_0(x) + F_k(u, h(x)) + c\sigma(h(x)). \tag{28}$$

The following simple lemma computes L over the set $\Lambda_0 := \{0\} \times \mathbb{N}$, and shows that the latter set is directed for L in the sense of Definition 14.

Lemma 22. *The following hold for the sequence $\{L_k\}$ defined in (28) and for the set $\Lambda_0 := \{0\} \times \mathbb{N}$.*

(i) Assume that h is continuous. Then for all $n \in \mathbb{N}$ and all $x \in \mathbb{R}^n$ we have

$$L(x, (0, n)) = f_0(x) + n\,\sigma(h(x)).$$

(ii) Λ_0 is directed for L.

Proof: (i) Definition 21 and (27) imply that, for all $(0, n) \in \Lambda_0$ and every $x \in X$ we have

$$\begin{aligned} L(x, (0, n)) &= \liminf_k L_k(x, (0, n)) \\ &= f_0(x) + n\sigma(h(x)) + \liminf_k F_k(0, h(x)) \\ &= f_0(x) + n\sigma(h(x)) = L_k(x, (0, n)), \end{aligned} \tag{29}$$

which proves (i). Statement (ii) follows as in Theorem 15. ♠

Proposition 23. Assume that f_0 is lower semicontinuous and h is continuous. Consider the set $\Lambda_0 := \{(0, n) : n \in \mathbb{N}\}$. The Lagrangian L and the sequence $\{L_k\}$ defined in (28) satisfy H_0 and $H_3(\Lambda_0)$. Condition $H_1(\Lambda_0)$ holds for L and the sequence $\{L_k\}$ if

(a) f_0 is bounded below, i.e., there exists $b_0 \in \mathbb{R}$ such that $f_0(x) \geqslant b_0$ for all $x \in X$,

(b) σ has a valley at 0, and

(c) a global error bound holds for h.

Proof: Let us first prove that $H_1(\Lambda_0)$ holds for L and the sequence $\{L_k\}$ under (a)–(c). By (29) we have

$$L(x,(0,n)) = f_0(x) + n\sigma(h(x)) = L_k(x,(0,n)),$$

so by Theorem 11 it is enough to show that $H_1(\Lambda_0)$ holds for L. Assume that (a)–(c) hold and $H_1(\Lambda_0, L)$ does not hold. This implies the existence of $\alpha_0 < M_P$, $\delta_0 > 0$, and a sequence $\{x_n\} \subset X$ with $d(x_n, X_0) \geqslant \delta_0$ such that

$$f_0(x_n) + n\sigma(h(x_n)) \leqslant \alpha_0. \tag{30}$$

We claim that the sequence $\{x_n\}$ cannot be bounded. Indeed, if there exists R_0 such that $\|x_n\| \leqslant R_0$ then there exists a subsequence of $\{x_n\}$, convergent to some point $\bar{x}$ such that $\|\bar{x}\| \leqslant R_0$ and $d(\bar{x}, X_0) \geqslant \delta_0$. Without loss of generality we still denote the convergent subsequence by $\{x_n\}$. Since $\bar{x} \notin X_0$ we must have that $\bar{r} := \|h(\bar{x})\| > 0$. Since h is continuous for n large enough we have that $\|h(x_n)\| > \bar{r}/2 > 0$. Using the lower semicontinuity of f_0 and taking limits in (30) we have

$$\alpha_0 \geqslant f_0(\bar{x}) + \liminf_n n\sigma(h(x_n)) \geqslant b_0 + c_{\bar{r}/2} \liminf_n n = +\infty,$$

where we used (a), and (b) with $\delta := \bar{r}/2$ in Definition 18. The above inequality entails a contradiction and hence we must have $\{x_n\}$ unbounded. Using now assumption (c) there exists $\gamma > 0$ such that

$$\|h(x_n)\| \geqslant \frac{d(x_n, X_0)}{\gamma} \geqslant \frac{\delta_0}{\gamma},$$

which, together with assumption (b) with $\delta := \delta_0/\gamma$ in Definition 18 yields

$$\sigma(h(x_n)) \geqslant c_{(\delta_0/\gamma)}.$$

Combining this inequality with assumption (a) and (30) again yields a contradiction. This proves that $H_1(\Lambda_0, L)$ holds under assumptions (a)–(c). Hence the statement on H_1 has been proved. We proceed now to show H_0 for L (and for L_k). For $x \in X_0$, we have that $h(x) = 0$ and by (26–27)

$$L(x, \lambda) = L_k(x, \lambda) = f_0(x) + F_k(u, 0) + c\sigma(0) = f_0(x),$$

so H_0 holds for L and all L_k. For every $(0, n) \in \Lambda_0$ and every $x \in X$ we have

$$\begin{aligned} f_0(x) &\leqslant f_0(x) + n\sigma(h(x)) \\ &= f_0(x) + F_k(0, h(x)) + n\sigma(h(x)) \\ &= L_k(x, (0, n)) \\ &= L(x, (0, n)), \end{aligned}$$

where we used that $F_k(0, h(x)) = 0$ for every $x \in X$, and that $\sigma(\cdot) \geqslant 0$. Hence,

$$f_0(x) \leqslant \liminf_k L_k(x, (0, n)) = L(x, (0, n)), \tag{31}$$

which implies that $H_3(\Lambda_0)$ holds for L and all L_k. ♠

The previous result readily implies strong asymptotic duality for (P).

Corollary 24. Assume that the hypotheses of Proposition 23 hold. If $H_2(L)$ holds, then the sequence $\{L_k\}$ defined in (28) verifies

$$\lim_k M_{D_k} = M_P = M_D.$$

Proof: Let $\Lambda_0 := \{(0, n) \; : \; n \in \mathbb{N}\}$. Proposition 23 shows that the sequence $\{L_k\}$ and L verify H_0. It also establishes H_1 and H_3 for L and L_k. It only remains to check that the set Λ_0 is directed for L. This is precisely Lemma 22(ii). The conclusion now follows from Corollary 17(ii). ♠

Remark 25. For the equality constrained problem as formulated in this section we can use a sequence of Lagrangians of the form

$$L_k(x, \lambda) := f_0(x) + \langle u, A_k h(x) \rangle + c\sigma(h(x)), \tag{32}$$

where $\{A_k\} \in \mathbb{R}^{m \times m}$ is a sequence of symmetric matrices. This is a particular case of Definition 21. As long as $H_2(L)$ and conditions (a)–(c) of Proposition 23 hold, the sequence of dual values will converge to the optimal primal value. This choice is of interest if we want to emphasize different constraints at each iteration. In this situation, the matrix A_k can be dynamically updated so that only a subset of the constraints is considered at the iteration k.

Remark 26. Condition $H_2(L)$ can be enforced, for instance, if we know that (P) has solutions in a given ball $B[0, R] := \{x \in \mathbb{R}^n \ : \ \|x\| \leqslant R\}$. In this case, in the formulation of (P) we can replace the constraint function h by $\hat{h}(x) := \max\{0, \|h(x)\|, \|x\| - R\}$, which gives a compact constraint set $\hat{X}_0 := \{x \ : \ \hat{h}(x) = 0\} \subset B[0, R]$. Then for every $r > 0$, $H_2(L)$ trivially holds because

$$X(r) = \hat{X}_0 + B[0, r],$$

which is compact as it is the Minkowski sum of two compact sets.

3.1 When f_0 is not bounded below

Proposition 23 cannot be applied when the objective function f_0 is linear, because assumption (a) is not true. The following result applies to the case of linear f_0. Recall that given a set $C \subset \mathbb{R}^m$, the *indicator function of the set* C is denoted by $\delta_C : \mathbb{R}^m \to \mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$ and defined as

$$\delta_C(z) = \begin{cases} 0 & \text{if } z \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

Given a function $g : \mathbb{R}^m \to \mathbb{R}_\infty$, recall that the *domain of* g is the set $\operatorname{dom} g := \{x \in \mathbb{R}^n \ : \ g(x) < +\infty\}$.

For the analysis of this case, we use the following sequence of Lagrangians. We assume that one feasible point is available, i.e., there is some $\hat{x} \in X_0$.

Definition 27. Assume $\hat{x} \in X_0$. Given $k \in \mathbb{N}$, denote by $B[0,k] := \{y \in \mathbb{R}^m : \|y\| \leqslant k\}$ the closed ball of center 0 and radius k in $\mathbb{R}^m$. Define $\sigma_k : \mathbb{R}^m \to \mathbb{R}_\infty$ as

$$\sigma_k(y) = \begin{cases} \sigma(y) & \text{if } \|y\| \leqslant k, \\ +\infty & \text{otherwise,} \end{cases} \tag{33}$$

where σ is as in Definition 18. In other words, $\sigma_k = \sigma + \delta_{B[0,k]}$, where $\delta_{B[0,k]}$ is the indicator function of the set $B[0,k]$. To define our sequence of Lagrangians, let $\Lambda := \mathbb{R}^m \times \mathbb{R}_+$. For each $k \in \mathbb{N}$ and each $\lambda = (u,c) \in \mathbb{R}^m \times \mathbb{R}_+$, let $L_k : \mathbb{R}^n \times \Lambda \to \mathbb{R}_\infty$ be defined as follows

$$L_k(x,\lambda) := f_0(x) + c\,\|x - \hat{x}\|\sigma_k(h(x)) + F_k(u, h(x)). \tag{34}$$

The following simple lemma computes L over the set $\Lambda_0 := \{0\} \times \mathbb{N}$, and shows that the latter set is directed for L in the sense of Definition 14.

Lemma 28. *The following hold for the sequence $\{L_k\}$ defined in (34) and for the set $\Lambda_0 := \{0\} \times \mathbb{N}$.*

(i) Assume that h is continuous. Then for all $n \in \mathbb{N}$ we have

$$L(x,(0,n)) = f_0(x) + n\,\|x - \hat{x}\|\,\sigma(h(x)).$$

(ii) Λ_0 is directed for L.

Proof: (i) Assume $x \in \mathbb{R}^n$ and $n \in \mathbb{N}$. Since h is continuous $\|h(x)\| \in \mathbb{R}$ and hence there exist (a unique) $j(x) \in \mathbb{N}$ such that

$$j(x) - 1 < \|h(x)\| \leqslant j(x).$$

Assume $p \in \mathbb{N}$ such that $p \geqslant j(x)$. Then $h(x) \in B[0,j(x)] \subset B[0,p]$ and hence $\sigma_p(h(x)) = \sigma(h(x)) + \delta_{B[0,p]}(h(x)) = \sigma(h(x))$. Hence, for $p \geqslant j(x)$ we have

$$\inf_{k\geqslant p} L_k(x,(0,n)) = f_0(x) + n\|x-\hat{x}\|\sigma_k(h(x)) = f_0(x) + n\|x-\hat{x}\|\sigma(h(x)).$$

If $p < j(x)$ then by definition of $j(x)$ we have that $h(x) \notin B[0,p]$ and hence $\sigma_p(h(x)) = +\infty$. Therefore, the infimum for $k \geqslant p$ will not take into account the values of k which are less than $j(x)$. Namely, for $p < j(x)$

$$\inf_{k\geqslant p} L_k(x,(0,n)) = \inf_{k\geqslant j(x)} L_k(x,(0,n)) = f_0(x) + n\|x-\hat{x}\|\sigma(h(x)).$$

Since in neither case the $\inf_{k\geqslant p} L_k(x,(0,n))$ depends on p, we have

$$L(x,(0,n)) = \sup_p \inf_{k\geqslant p} L_k(x,(0,n)) = f_0(x) + n\|x-\hat{x}\|\sigma(h(x)).$$

For proving (ii), assume $n, m \in \mathbb{N}$. Let $p := \max\{n,m\}$. Then by part (i)

$$\begin{aligned}
\max\{L(x,(0,n)), L(x,(0,m))\} &= f_0(x) + \|x-\hat{x}\|\sigma(h(x))\max\{n,m\} \\
&= f_0(x) + p\|x-\hat{x}\|\sigma(h(x)) \\
&= L(x,(0,p)).
\end{aligned}$$

♠

Proposition 29. Assume that f_0 is lower semicontinuous and h is continuous. Consider the set $\Lambda_0 := \{(0,n) \,:\, n \in \mathbb{N}\}$. Then H_0 holds for L and L_k for all k. Also $H_3(\Lambda_0, L)$ holds. Moreover, condition $H_1(\Lambda_0)$ holds for L and the sequence $\{L_k\}$ if

(a) $\liminf_{\|x\|\to\infty} \dfrac{f_0(x)}{\|x\|} = b_0 \in \mathbb{R}$,

(b) A global error bound holds for h,

(c) There exists $a > 0$ such that $\sigma(z) \geqslant a\|z\|$ for all $z \in \mathbb{R}^m$.

Proof: We start by proving H_0 for L and L_k. Since $\sigma_k(0) = 0$, for $x \in X_0 \cap \operatorname{dom} f_0$ by (26) and (27)

$$L_k(x,(u,c)) = f_0(x) + F_k(u,0) + c\|x-\hat{x}\|\sigma_k(0) = f_0(x) \in \mathbb{R}.$$

If $x \in X_0 \setminus \operatorname{dom} f_0$ then by definition of L_k we have $L_k(x,\lambda) = f_0(x) = +\infty$. In all cases, $L_k(x,\lambda) = f_0(x)$ for all $x \in X_0$, $\lambda \in \Lambda$. So H_0 holds for all L_k. Using the latter equality and the definition of L,

$$L(x,\lambda) = \liminf_k L_k(x,\lambda) := f_0(x), \text{ for all } x \in X_0,\ \lambda \in \Lambda,$$

so H_0 also holds for L. For checking H_3, we use Lemma 28(i) and that $\sigma \geqslant 0$ to write

$$L(x,(0,n)) = f_0(x) + n\|x-\hat{x}\|\sigma(h(x)) \geqslant f_0(x),$$

for all $x \in \mathbb{R}^n$. This shows that $H_3(\Lambda_0, L)$ holds. By (33), $\sigma_k \geqslant \sigma$, therefore

$$L_k(x,\cdot) \geqslant L(x,\cdot), \text{ over the set } \Lambda_0. \tag{35}$$

Hence, by Theorem 11, if $H_1(\Lambda_0)$ holds for L, it will hold for all L_k's. Let us then check $H_1(\Lambda_0, L)$ under (a)–(c). By Lemma 28(i) for all $(0,n) \in \Lambda_0$ and every $x \in X$

$$L(x,(0,n)) = f_0(x) + n\|x-\hat{x}\|\sigma(h(x)). \tag{36}$$

Assume that $H_1(\Lambda_0, L)$ does not hold. This implies the existence of $\alpha_0 < M_P$, $\delta_0 > 0$, and a sequence $\{x_n\} \subset X$ with $d(x_n, X_0) \geqslant \delta_0$ such that

$$f_0(x_n) + n\|x_n - \hat{x}\|\sigma(h(x_n)) \leqslant \alpha_0. \tag{37}$$

As in the proof of Proposition 23, we claim that the sequence $\{x_n\}$ is unbounded. Indeed, assume the sequence is bounded. Then it has a convergent subsequence, which by simplicity we still denote as $\{x_n\}$. Denote by $\bar{x}$ the limit of this sequence. Using (b), (c), and the lower semicontinuity of f_0

$$\begin{aligned}
\alpha_0 &\geqslant \liminf_n \,[f_0(x_n) + n\,\|x_n - \hat{x}\|\sigma(h(x_n))] \\
&\geqslant \liminf_n \,[f_0(x_n) + n\,a\,\|x_n - \hat{x}\|\,\|h(x_n)\|] \\
&\geqslant \liminf_n \,[f_0(x_n) + n\,a\,\|x_n - \hat{x}\|\,\gamma\,d(x_n, X_0)] \\
&\geqslant \liminf_n \,\left[f_0(x_n) + n\,a\gamma\,{\delta_0}^2\right] \\
&\geqslant f_0(\bar{x}) + \,a\gamma{\delta_0}^2 \liminf_n \,n \\
&= +\infty,
\end{aligned}$$

where we also used that $\|x_n - \hat{x}\| \geqslant d(x_n, X_0) \geqslant \delta_0$. The above expression entails a contradiction. Hence our claim is true and $\{x_n\}$ is unbounded. We can assume that the whole sequence $\{\|x_n\|\}$ tends to infinity. So the sequence $\{a_n\} \subset \mathbb{R}$ defined as

$$a_n := \frac{\|x_n\|}{\|x_n - \hat{x}\|} = \frac{1}{\dfrac{\|x_n - \hat{x}\|}{\|x_n\|}} = \frac{1}{\left\|\dfrac{x_n}{\|x_n\|} - \dfrac{\hat{x}}{\|x_n\|}\right\|},$$

tends to 1 when $n \to \infty$. Hence, there exists n_0 such that for all $n \geqslant n_0$ we have $a_n < 2$. Using (37) and assumptions (a), (b) and (c) we can write, for $n \geqslant n_0$,

$$\begin{aligned}
\frac{\alpha_0}{\|x_n - \hat{x}\|} &\geqslant \frac{f_0(x_n)}{\|x_n - \hat{x}\|} + n\,\sigma(h(x_n)) \\
&\geqslant \frac{f_0(x_n)}{\|x_n\|}\frac{\|x_n\|}{\|x_n - \hat{x}\|} + n\,a\,\|h(x_n)\| \\
&\geqslant b_0\,a_n + n\,a\,\gamma d(x_n, X_0) \\
&\geqslant -|b_0|\,a_n + n\,a\,\gamma\delta_0 \\
&\geqslant \,-2\,|b_0| + n\,a\,\gamma\delta_0,
\end{aligned}$$

which is a contradiction because the left hand side tends to zero while the right hand side tends to infinity. This proves that $H_1(\Lambda_0, L)$ holds under assumptions (a)–(c). This, together with (35) and Theorem 11, show that H_1 also holds for L_k. This completes the proof. ♠

The previous proposition readily implies strong asymptotic duality for (P) for the case in which f_0 is not bounded below.

Corollary 30. Assume that the hypotheses of Proposition 29 hold. If $H_2(L)$ holds, then the sequence $\{L_k\}$ defined in (34) verifies

$$\lim_k M_{D_k} = M_P = M_D.$$

Proof: The conclusion holds because, by Proposition 29, all assumptions of Corollary 17(ii) hold for our sequence of Lagrangians. Indeed, Proposition 29 shows that H_0 holds for L_k and L. It also establishes H_1 and H_3 for L. Condition H_1 holds for all L_k by (35) and Theorem 11. The only remaining fact to check is that the set Λ_0 is directed for L. This is precisely Lemma 28(ii). The conclusion now follows from Corollary 2.2(ii). ♠

4 Conclusions

We have presented a duality approach based on a sequence of dual problems, which, under certain basic assumptions, have optimal values converging to the optimal primal value. This is a first step towards the development of primal-dual schemes that admit a dynamic update of the Lagrangian function. We reviewed the existing literature on asymptotic duality and presented specific cases in which strong duality holds in the asymptotic sense. Asymptotic properties can be useful for approximating optimal values or solutions of (P) by using a suitable sequence of Lagrangians. Our future research will use the theory presented here for solving specific families of nonconvex problems (e.g., linear integer programming problems, polynomial optimization). For general nonconvex problems, good asymptotic properties can help in providing new ways for approximating the solution. Namely, new convergence results can be established for primal-dual techniques in which the search direction used for solving the dual problem is updated by using a suitable sequence of Lagrangians $\{\mathsf{L}_k\}$.

References

[1] R. S. Burachik, W. P. Freire and C. Y. Kaya. Interior epigraph directions method for nonsmooth and nonconvex optimization via generalized augmented Lagrangian duality. *Journal of Global Optimization*, **60(3)**:501–529, 2014. doi:10.1007/s10898-013-0108-4

[2] R. S. Burachik, R. N. Gasimov, N. A. Ismayilova and C. Y. Kaya. On a modified subgradient algorithm for dual problems via sharp augmented Lagrangian. *Journal of Global Optimization*, **34(1)**:55–78, 2006. doi:10.1007/s10898-005-3270-5 C95

[3] R. S. Burachik, A. N. Iusem and J. G. Melo. The exact penalty map for nonsmooth and nonconvex optimization. *Optimization*, **64(4)**:717–738, 2015. doi:10.1080/02331934.2013.830117 C95

[4] R. S. Burachik, A. N. Iusem and J. G. Melo. An inexact modified subgradient algorithm for primal-dual problems via augmented Lagrangians. *Journal of Optimization Theory and Applications*, **157(1)**:108–131, 2013. doi:10.1007/s10957-012-0158-7 C95

[5] R. S. Burachik, A. N. Iusem and J. G. Melo. A primal dual modified subgradient algorithm with sharp Lagrangian. *Journal of Global Optimization*, **46(3)**: 347–361, 2010. doi:10.1007/s10898-009-9429-8 C95

[6] R. S. Burachik, A. N. Iusem and J. G. Melo. Duality and exact penalization for general augmented Lagrangians, *Journal of*

Optimization Theory and Applications, **147(1)**:125–140, 2010. doi:10.1007/s10957-010-9711-4 C95

[7] R. S. Burachik and C. Y. Kaya. An augmented penalty function method with penalty parameter updates for nonconvex optimization, *Nonlinear Analysis: Theory, Methods & Applications*, **75(3)**:1158–1167, 2012. doi:10.1016/j.na.2011.03.013 C95

[8] R. S. Burachik and C. Y. Kaya. An update rule and a convergence result for a penalty function method. *Journal of Industrial Management and Optimization*, **3**(2): 381–398, 2007. doi:10.3934/jimo.2007.3.381 C95

[9] R. S. Burachik, C. Y. Kaya and M. Mammadov. An inexact modified subgradient algorithm for nonconvex optimization. *Computational Optimization and Applications*, **45(1)**:1–24, 2010. doi:10.1007/s10589-008-9168-7 C95

[10] R. S. Burachik and X. Q. Yang, Asymptotic strong duality, *Numerical Algebra, Control and Optimization (NACO)*, **1(3)**:539–548, 2011. doi:10.3934/naco.2011.1.539 C96, C100, C101, C102, C103, C104, C105, C106

[11] R. S. Burachik and A. M. Rubinov. Abstract convexity and augmented Lagrangians. *SIAM Journal on Optimization*, **18(2)**:413–436, 2007. doi:10.1137/050647621 C95, C111

[12] R. S. Burachik and A. M. Rubinov, On the absence of duality gap for Lagrange-type functions, *Journal of Industrial and Management Optimization*, **1(1)**:33–38, 2005. doi:10.3934/jimo.2005.1.33 C102, C103

[13] R. N. Gasimov. Augmented Lagrangian duality and nondifferentiable optimization methods in nonconvex programming. *Journal of Global Optimization*, **24(2)**:187–203, 2002. doi:10.1023/A:1020261001771 C95

[14] R. N. Gasimov and A. M. Rubinov. On augmented Lagrangians for optimization problems with a single constraint. *Journal of Global*

Optimization, **28(2)**:153–173, 2004.
doi:10.1023/B:JOGO.0000015309.88480.2b C95

[15] O. Güler. *Foundations of Optimization*, Graduate texts in Mathematics 258, Springer, Berlin, 2010. doi:10.1007/978-0-387-68407-9 C111

[16] G. Li. Global error bounds for piecewise convex polynomials. *Mathematical Programming*, **137(1)**:37–64, 2013. doi:10.1007/s10107-011-0481-z C111

[17] X. X. Huang, K. L. Teo and X. Q. Yang, Approximate augmented Lagrangian functions and nonlinear semidefinite programs, *Acta Mathematica Sinica, English Series*, **22**:1283–1296, 2006. doi:10.1007/s10114-005-0702-6 C95, C96

[18] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998. doi:10.1007/978-3-642-02431-3 C95, C106

[19] A. M. Rubinov, X. X. Huang and X. Q. Yang, The zero duality gap property and lower semicontinuity of the perturbation function, *Mathematics of Operations Research*, **27**:775–791, 2002. doi:10.1287/moor.27.4.775.295 C95

[20] C. Y. Wang, X. Q. Yang and X. M. Yang, A unified nonlinear Lagrangian approach to duality and optimal paths, *Journal of Optimization Theory and Applications* **135**:85–100, 2007. doi:10.1007/s10957-007-9225-x C95, C111

Author address

1. **Regina S. Burachik**, School of Information Technology and Mathematical Sciences, University of South Australia, Australia.
`mailto:regina.burachik@unisa.edu.au`